Final Project in Causal Inference

Rotem Torkan 212705313; Maor Dikter 214169377

February 2023

How does college students' displacement affect their academic success?

Abstract

This project in causal inference aims to investigate the relationship between college students' displacement and their academic success. This study seeks to find the average treatment effect of displacement on academic success, using various statistical methods learnt in the course. We believe that the findings of this study will provide important insights into the effects of displacement on academic outcomes for college students, and can inform policies and interventions aimed at supporting these students.

Contents

1 Introduction

1.1 About Our Dataset

Our dataset contains data from a higher education institution in Portugal on various variables. Our target values are whether the student graduated, dropped out, or still enrolled at the time.

This dataset provides a comprehensive view of students enrolled in various undergraduate degrees. It includes demographic data, social-economic factors and academic performance information that can be used to analyze the possible predictors of student dropout and academic success. This dataset contains relevant information available at the time of enrollment, such as application mode, marital status, course chosen and more.

Our dataset also provides economic information about unemployment rate, inflation rate and GDP from the region which can help us further understand how economic factors play into student dropout rates or academic success outcomes.

1.2 Our Causal Question

The question we'll try to answer is "How does college students' displacement affect their academic success?"

A student being displaced means that he is no longer living in their usual home. This could affect their academic success both positively and negatively. In our dataset, academic success is measured in 3 potential outcomes; Graduated, if the student had completed his studies successfully. Dropout, if the student had failed and dropped out, and enrolled, if the student was still registered after the period of time allocated for studying, meaning that he didn't fulfill his requirements on time. We consider binary academic success, where student is successful if and only if she graduated on time. Mathematically,

$$\mathbb{1}[successful] = \begin{cases} 1 & \text{If Target == Graduate} \\ 0 & \text{Otherwise, if Target == Dropout / Enrolled} \end{cases}.$$

In our project, we plan to understand how the displacement of a student affects their target outcome by estimating various types of affects e.g., IPW, S-Learner and matching. One of the challenges for this project could be the existance of hidden confounders such as siblings historic academic success, which does not appear in the data.

2 The Data

The data refers to records of students enrolled between the academic years 2008/2009 to 2018/2019. These include data from 17 undergraduate degrees from different fields of knowledge, such as agronomy, design, education, nursing, journalism, management, social service, and technologies. It contains 4424 records, where each record represents an individual student.

2.1 Treatment, Features and Potential Outcomes

The dataset presented above contains 36 different features and 1 outcome (a total of 37 columns). It can be found here: Predict Dropout or Academic Success | Kaggle

The features included in our data are from various fields; demographic data, socioeconomic and macroeconomic data, data at the time of student enrollment, and data at the end of the first and second semesters. As mentioned, we chose the treatment feature to be the displacement of the student, meaning this project tries to find the effect of a demographic feature on academic success.

The features divided into classes:

Class of Attribute	Attribute	Type
Demographic data	Marital status	Numeric/discrete
	Nationality	Numeric/discrete
	Displaced	Numeric/binary
	Gender	Numeric/binary
	Age at enrollment	Numeric/discrete
	International	Numeric/binary
Socioeconomic data	Mother's qualification	Numeric/discrete
	Father's qualification	Numeric/discrete
	Mother's occupation	Numeric/discrete
	Father's occupation	Numeric/discrete
	Educational special needs	Numeric/binary
	Debtor	Numeric/binary
	Tuition fees up to date	Numeric/binary
	Scholarship holder	Numeric/binary
Macroeconomic data	Unemployment rate	Numeric/continuous
	Inflation rate	Numeric/continuous
	GDP	Numeric/continuous
Academic data at enrollment	Application mode	Numeric/discrete
	Application order	Numeric/ordinal
	Course	Numeric/discrete
	Daytime/evening attendance	Numeric/binary
	Previous qualification	Numeric/discrete
Academic data at the end of 1st semester	Curricular units 1st sem (credited)	Numeric/discrete
	Curricular units 1st sem (enrolled)	Numeric/discrete
	Curricular units 1st sem (evaluations)	Numeric/discrete
	Curricular units 1st sem (approved)	Numeric/discrete
	Curricular units 1st sem (grade)	Numeric/continuous
	Curricular units 1st sem (without evaluations)	Numeric/discrete
Academic data at the end of 2nd semester	Curricular units 2nd sem (credited)	Numeric/discrete
	Curricular units 2nd sem (enrolled)	Numeric/discrete
	Curricular units 2nd sem (evaluations)	Numeric/discrete
	Curricular units 2nd sem (approved)	Numeric/discrete
	Curricular units 2nd sem (grade)	Numeric/continuous
	Curricular units 2nd sem (without evaluations)	Numeric/discrete
Target	Target	Categorical

More details about every feature:

- **Marital status:** The marital status of the student. Can be one of the following: 1—Single 2—Married 3—Widower 4—Divorced 5—Facto union 6—Legally separated.
- **Nationality:** The nationality of the student. Possible values are from 1 to 21, 1—Portuguese, 2 to 21—Other.
- **Displaced:** A student that is no longer living in their usual home, 0—Settled, 1—Displaced.
- **Gender:** The gender of the student, 0—female, 1—male.
- **Age at enrollment:** The age of the student at enrollment. Values are from 17 to 70.
- **International:** Whether the student is international or not.
- **Previous qualification:** A student's previous level of education. Values are from 1 to 17, each referring to a different level of education.
- **Previous qualification (grade):** The grade of the previous qualification. Values are from 0 to 200.
- **Mother's/Father's qualification:** The level of education of the student parent. Values are from 1 to 34, each referring to a different level of education.
- **Mother's/Father's occupation:** The current occupation of the student parent. Values are from 1 to 46, each referring to a different field of employment.
- **Educational special needs:** Whether the student requires educational special needs. 0—No, 1—Yes.
- **Debtor:** Whether a student has a debt, 0—No, 1—Yes.

- **Tuition fees up to date:** Whether the student has paid his tuition fees on time, 0—No, 1—Yes.
- **Scholarship holder:** Whether the student received a scholarship, 0—No, 1—Yes.
- **Unemployment rate:** The average unemployment rate in the student's study years, continuous values.
- **Inflation rate:** The average inflation rate in the student's study years, continuous values.
- **GDP:** The average GDP in the student's study years, continuous values.
- **Application mode:** The type of the student's application. Values are from 1 to 18.
- **Application order:** The student's preference (of the application). Values are from 1 to 9, when 1 is the highest and 9 is the lowest.
- **Course:** Field of study. Values are from 1 to 17.
- **Daytime/evening attendance:** Whether the student studies during the daytime, 0—No, 1—Yes.
- **Curricular units 1st/2nd sem:** The student's credit at the 1st/2nd semester. There are 6 different features for each semester: credited, enrolled, evaluations, approved, grade and without evaluations.

3 Preprocessing and Analysis

3.1 Data Preprocessing

The data was comfortable to work with, since most of the features have numerical values. We loaded our CSV file into a Pandas dataframe and worked with it. The main challenge in the processing of our data was to determine which features to keep as confounders and which features to drop before calculating the ATE. The columns we dropped and were not taken in the calculations were:

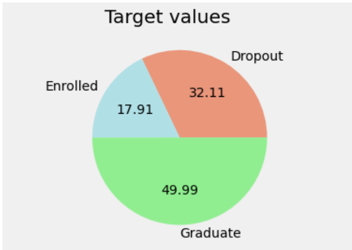
Features related to the curricular units: *"Curricular units 1st sem (credited)", "Curricular units 1st sem (enrolled)", "Curricular units 1st sem (evaluations)", "Curricular units 1st sem (approved)", "Curricular units 1st sem (grade)", "Curricular units 1st sem (without evaluations)", "Curricular units 2nd sem (credited)", "Curricular units 2nd sem (enrolled)", "Curricular units 2nd sem (evaluations)", "Curricular units 2nd sem (approved)", "Curricular units 2nd sem (grade)", "Curricular units 2nd sem (without evaluations)"*

And the feature for whether a student require educational special needs, *"Educational special needs"*.

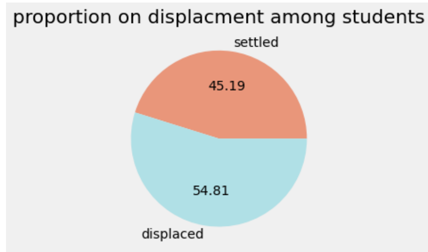
We decided to drop any feature related to the curricular units since it is directly related to the target. For finding an unbiased relation between our target and a feature known at the time of enrollment we cannot use any feature given from a later time then enrollment. We decided to drop the "Educational special needs" feature since we didn't have sufficient data about the students requiring educational special needs (only 50 students). In addition, the proportion of displacement among students with educational special needs is approximately balanced (29 displaced / 22 not), considering the small number of examples we have.

3.2 Data Analysis

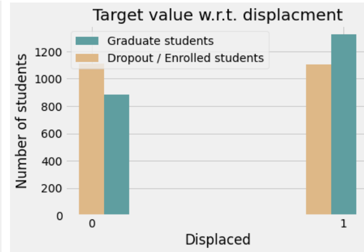
In order to learn about our data we first plotted the proportion of the target values among the students [Figure 1a], the proportion of our treatment among the students [Figure 1b], and the proportion of the target among the students with respect to the treatment [Figure 1c].



(a) Proportion of target values



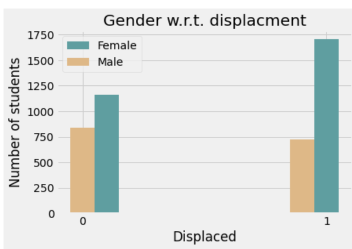
(b) Proportion of displacement among students



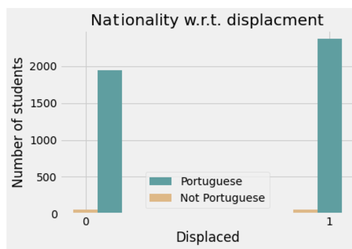
(c) Target value w.r.t displacement

We can see that 49.9% of the students graduated. Considering enrolled and dropout as the same case causes the target values to split evenly. We can also see that the treatment proportion is relatively balanced among the students.

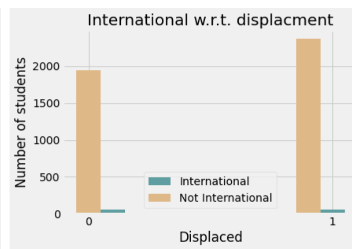
To learn more about our data and to decide which features to keep, we plotted the different features with respect to displacement. Binary features: gender [Figure 2a], nationality [Figure 2b], international [Figure 2c], age (as a binary variable- above or under 22 years old) [Figure 2d], previous qualification grade (as a binary variable- above ar under the median grade) [Figure 2e], daytime/evening attendance [Figure 2f], debtor [Figure 2g], tuition fees up to date [Figure 2h], scholarship holder [Figure 2i].



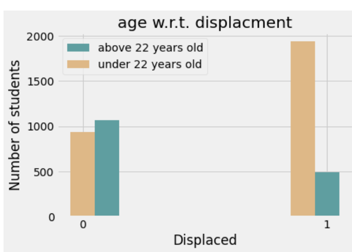
(a) Gender w.r.t displacement



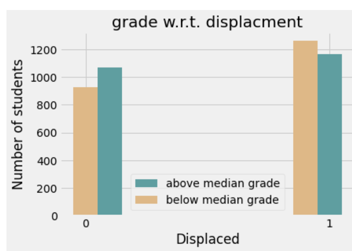
(b) Nationality w.r.t displacement



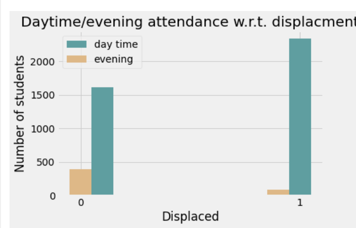
(c) International w.r.t displacement



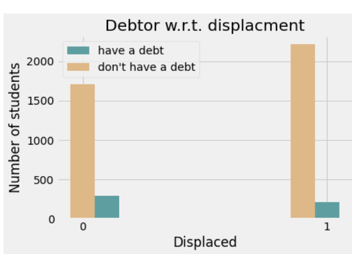
(d) Age w.r.t displacement



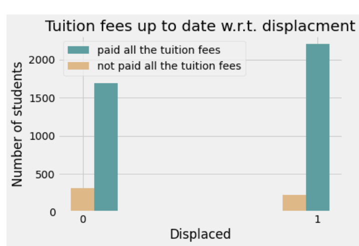
(e) Previous qualification (grade) w.r.t displacement



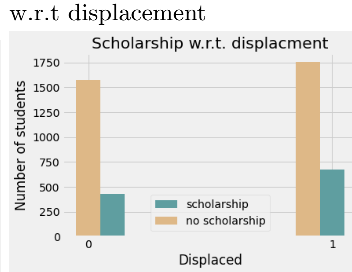
(f) Daytime/evening attendance w.r.t displacement



(g) Debtor w.r.t displacement

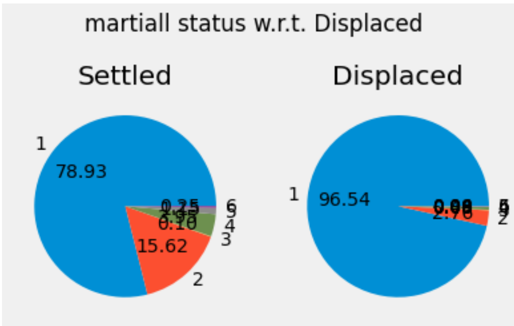


(h) Tuition fees up to date w.r.t displacement

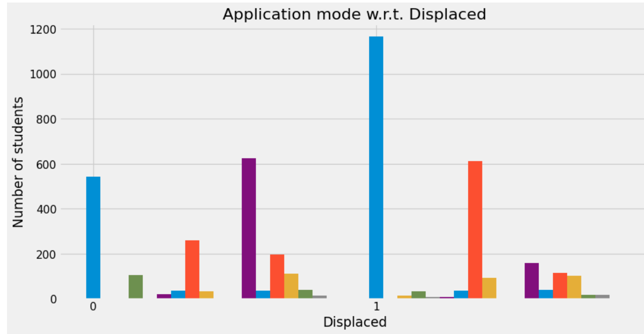


(i) Scholarship holder w.r.t displacement

Non-binary features: Marital status [Figure 3a], Application mode [Figure 3b], GDP [Figure 4].



(a) Marital Status w.r.t displacement



(b) Application mode w.r.t displacement

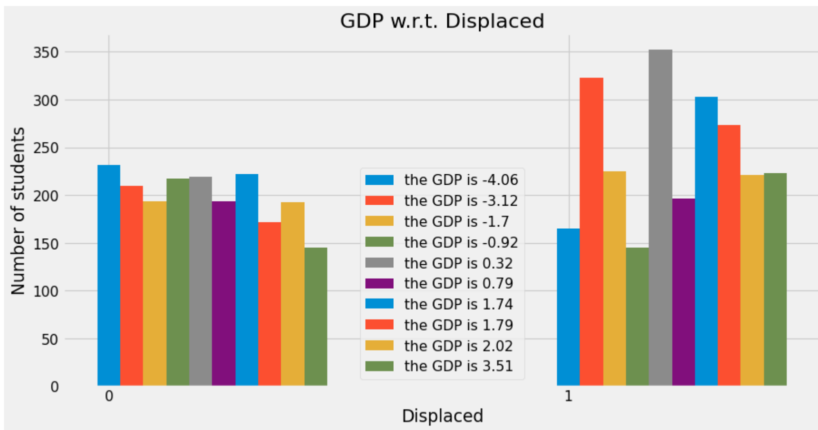


Figure 4: GDP w.r.t displacement

4 Assumptions and Possible Weaknesses

We have seen in class that in order to properly prove a causal inference claim, we must make sure that our data has several attributes.

These attributes are assumptions that ensure us the outcomes of our calculations are not related to some other parameters, that we did not take into account, so we can truly answer our causal inference question. The assumptions are: Stable Unit Treatment Value Assumption, consistency, ignorability and common support. In the next subsections, we address each assumption and explain why we can assume it in this work.

4.1 Stable Unit Treatment Value Assumption (SUTVA)

The SUTVA assumption refers to the following two assumptions:

1. The potential outcomes for any unit do not vary with the treatments assigned to other units.
2. For each unit, there are no different forms or versions of each treatment level, which lead to different potential outcomes.

We claim that the first assumption holds for our dataset since the fact that one student is displaced or settled does not affect the academic success of other students. As students, we know a lot of displaced and settled students and we can attest that other students displacement status does not affect the academic success.

The second assumption is a weakness of this project since it holds partially. On the one hand, one can argue that all the displacements are similar - the student left her original home and that's it. However, on the other hand, one can claim that the transition-distance matter and may affect the academic success. We believe that the distance effect is not significant and therefore we claim that the second assumption also holds.

4.2 Consistency

In order to hold the consistency assumption, for every unit that receives the treatment T, we should observe the corresponding potential outcome. Namely, $Y = TY_1 + (1-T)Y_0$. We claim that our study satisfies this assumption. The data was created from datasets of high education institutions. Since the dataset belongs to the universities, they can report the student academic status. In addition, the data was taken from the Ministry of Interior, which keeps the correct records and statistics about all the people. Therefore, we conclude that the data is correct and trustworthy, specifically the displacement status of every student (the treatment).

4.3 Ignorability

The ignorability assumption means that the potential outcomes are independent of treatment assignment, conditioned on observed covariates, namely, $(Y_0, Y_1) \perp\!\!\!\perp T|X$. In other words, ignorability holds if all existing confounders are measured. This assumption not necessary holds in our work, since the existence of unmeasured confounders is optional; in theory, there can be many hidden confounders, which have an indirect affect on the treatment and on the outcome. We think that it is fair to assume ignorability in this work since we addressed a lot of relevant confounders.

4.4 Common Support

The common support assumption holds if every set of features could belong to a subject from the treated or the untreated group. Formally, if $\forall t, x : P(T = t|X = x) > 0$. In theory, every set of features could belong to a settled student and to a displaced student. For example, even an international student can be settled as we show in [Figure 2c]. However, since the dataset is finite and there are a lot of features, there are feature sets that do not have a treated and/or untreated sample. We think that it is fair to assume the common support assumption in our case because it holds for most of the feature sets.

5 Treatment Effect Estimation

5.1 ATE Estimation

The Average Treatment Effect (ATE) is defined to be:

$$ATE = E[Y_1 - Y_0] = E[Y_1] - E[Y_0]$$

Calculating the ATE is an important part of our study and helps us determine the impact of the treatment on our outcome. In our case, the ATE is the difference between the expected value of graduating if treated, meaning being displaced, and the expected value of graduating if not treated, meaning staying settled.

There are various methods to calculate ATEs. In our study, we estimated the Average Treatment Effect using the methods inverse probability weighting (IPW), S-learner, T-learner, and matching. Each method has its own strengths and limitations.

5.1.1 Inverse Probability Weighting (IPW)

Inverse Probability Weighting (IPW) is a statistical method that estimates the ATE using the propensity scores- the weights of each observation in the dataset based on the probability of receiving the treatment.

The formula of the ATE using IPW is:

$$\widehat{ATE}_{IPW} = \frac{1}{n}\sum_{i=1}^{n}\frac{t_i y_i}{e(x_i)} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-t_i)y_i}{1-e(x_i)}$$

for $e(x) = p(T = 1|X = x)$ the propensity scores.

5.1.2 S-Learner

S-Learner is a machine learning method that estimates the treatment effect as a function of the covariates in the dataset.

In S-Learner, we fit a model $\hat{y} \approx f(x, t)$ with t as a feature on the entire sample. The formula of the ATE is given by:

$$\widehat{ATE}_{S-learn} = \frac{1}{n}\sum_{i=1}^{n}(f(x_i, 1) - f(x_i, 0))$$

5.1.3 T-Learner

T-Learner is also a method that estimates the treatment effect, but it differs from S-Learner in that it models the difference in outcomes between the treated and control groups directly.

In T-Learner we fit two separate models $\hat{Y}_1 \approx f_1(x), \hat{Y}_0 \approx f_0(x)$ on treated and control samples. The formula of the ATE is given by:

$$\widehat{ATE}_{T-learn} = \frac{1}{n}\sum_{i=1}^{n}(f_1(x_i) - f_0(x_i))$$

5.1.4 Matching

Matching is a method that creates pairs of observations that are similar in terms of their covariates and treatment status. In matching we assign for each sample x with treatment t, a closest sample $\tilde{x}$ with treatment 1-t. Using 1-NN matching, we estimate the ATE using an estimation of the Individual Treatment Effect (ITE) for each sample in the data, $\widehat{ITE}(i) = y_i - y_{j(i)}$, $j(i)$ is the nearest counterfactual neighbor of i. Hence, the formula of the ATE is given by:

$$\widehat{ATE}_{Matching} = \frac{1}{n}\sum_{i=1}^{n}\widehat{ITE}(i)$$

5.2 ATT Estimation

The Average Treatment Effect on the Treated (ATT) is defined to be:

$$ATT = E[Y_1 - Y_0|T = 1] = E[Y_1|T = 1] - E[Y_0|T = 1]$$

The ATT only includes the individuals who received the treatment.

5.2.1 Inverse Probability Weighting (IPW)

Similarly to the ATE, we estimates the ATT using the propensity scores. The formula of the ATT using IPW is:

$$\widehat{ATT}_{IPW} = \frac{\sum_{i=1}^{n} t_i y_i}{\sum_{i=1}^{n} T_i} - \frac{\sum_{i=1}^{n}(1-t_i)y_i \cdot \frac{e(x_i)}{1-e(x_i)}}{\sum_{i=1}^{n}(1-t_i) \cdot \frac{e(x_i)}{1-e(x_i)}}$$

for $e(x) = p(T = 1|X = x)$ the propensity scores.

5.2.2 S-Learner, T-Learner and Matching

The ATT is calculated with S-Learner, T-Learner or Matching exactly the same as the ATE is calculated with these methods (recall Subsection 5.1). The only difference is that we take the average only on the treated group in this case. Namely, for the S-Learner

$$\widehat{ATT}_{S-learn} = \frac{1}{|treated|}\sum_{x_i \in treated}(f(x_i, 1) - f(x_i, 0)),$$

where *treated* represents the treated group. The same holds for the T-Learner and the matching methods. Formally,

$$\widehat{ATT}_{T-learn} = \frac{1}{|treated|} \sum_{x_i \in treated} (f_1(x_i) - f_0(x_i))$$

and

$$\widehat{ATT}_{Matching} = \frac{1}{|treated|} \sum_{x_i \in treated} \widehat{ITE}(i).$$

5.3 Results

We calculated the ATEs and ATTs using the methods mentioned above. For the S-Learner and T-Learner methods we used 2 different learning algorithms- Decision Trees and Random Forest, that defined the model we wish to fit. More information and implementation can be found in our code. The results are presented in Table 1 and in bar plots; ATE scores in [Figure 5], ATT scores in [Figure 6].

Method	ATE	ATT
IPW	-0.002826	0.049462
S-learner (Decision Tree classifier)	0.004573	0.008343
S-learner (Random Forest classifier)	0.005030	0.009178
T-learner (Decision Tree classifier)	0.011205	0.043804
T-learner (Random Forest classifier)	0.030413	0.080517
Matching	0.018294	0.017521

Table 1: ATE and ATT scores

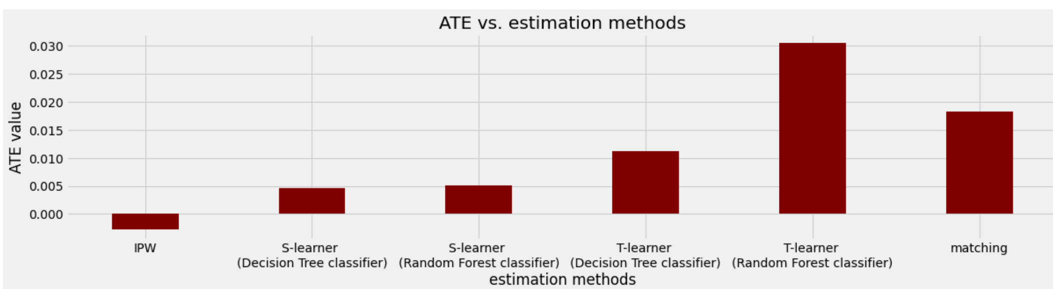


Figure 5: ATE scores for different estimation methods

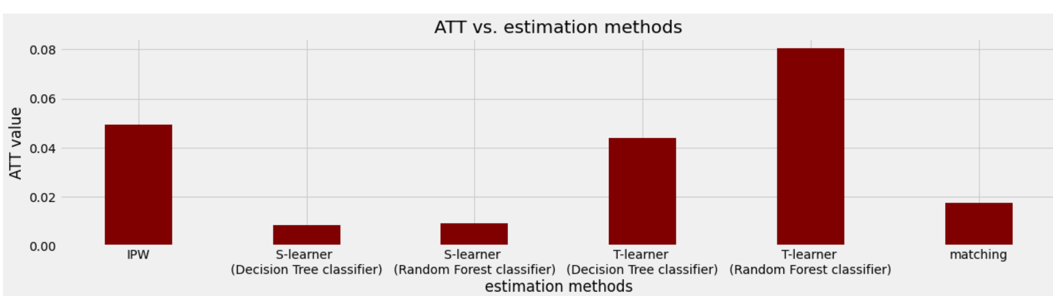
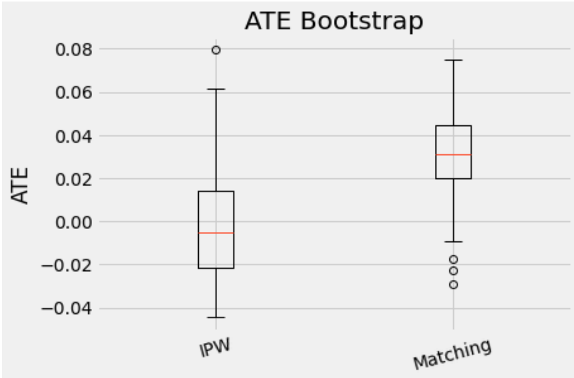
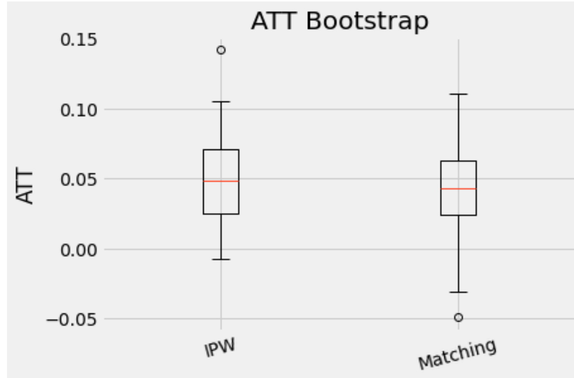


Figure 6: ATT scores for different estimation methods

To gain more precise probabilities, we used bootstrap on 2 of our methods (IPW and matching) to get a range of probabilities (a confidence interval) which gives us more precise results and avoids mistakes caused by high variance. We visualized the results with box-plots; ATE range in [Figure 7a], ATT range in [Figure 7b].



(a) Box-plot of the confidence interval of the ATE scores



(b) Box-plot of the confidence interval of the ATT scores

6 Conclusion and Discussion

Considering all of our methods, we can see that the ATE and ATT values are almost always positive and close to 0. Hence, we can understand from our results that the probability for a displaced student to graduate is slightly higher than the probability for a settled student to graduate.

In conclusion, this project has explored the relationship between college students' displacement and their academic success. Through the use of causal inference techniques, we have shown that displacement has a small positive impact on academic success.

While there is still much work to be done, this study provides a foundation for future research and intervention efforts aimed at improving academic studies and helping students with different characteristics.